# Data-driven identification of environmental variables influencing phenotypic plasticity for grain yield in hybrid maize

Aaron Kusmec[1], Dan Nettleton[2], Genomes to Fields Consortium, and Patrick S. Schnable[1,3]

Departments of [1]Agronomy and [2]Statistics and [3]Plant Sciences Institute,
Iowa State University, Ames, IA 50014, USA

INTRODUCTION

Phenotypes emerge over the course of development as the consequence of complex interactions between genotype and environment. Plant breeders seek to manipulate these interactions to produce cultivars with both high and stable yields. However, global climate change poses challenges to these objectives by shifting the average climate and increasing climate variability. To better meet these challenges, breeders require a better understanding of the environmental variables that underlie phenotypic plasticity. Current approaches include the use of process-based models leveraging physiological knowledge or exhaustive search of a candidate set of variables. These are limited because they either assume that environmental drivers will remain constant (process-based) or suffer from rapid expansion of the search space (exhaustive search). Here, we demonstrate the use of genetic algorithms to identify environmental variables underlying hybrid maize grain yield in the Genomes to Fields experiments and their utility in genetic analyses to generate hypotheses for the further study of plastic responses.
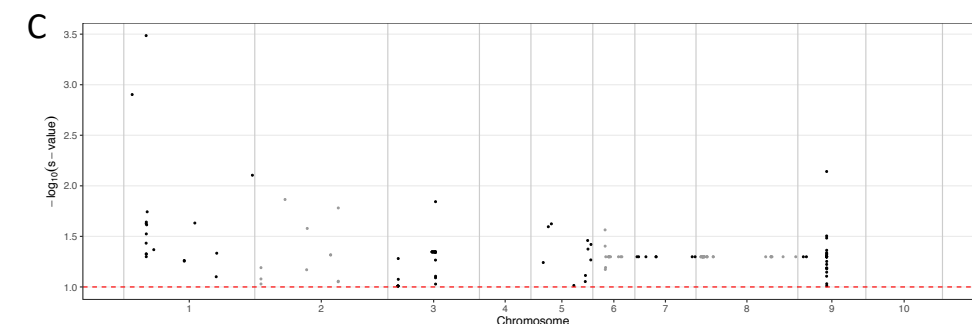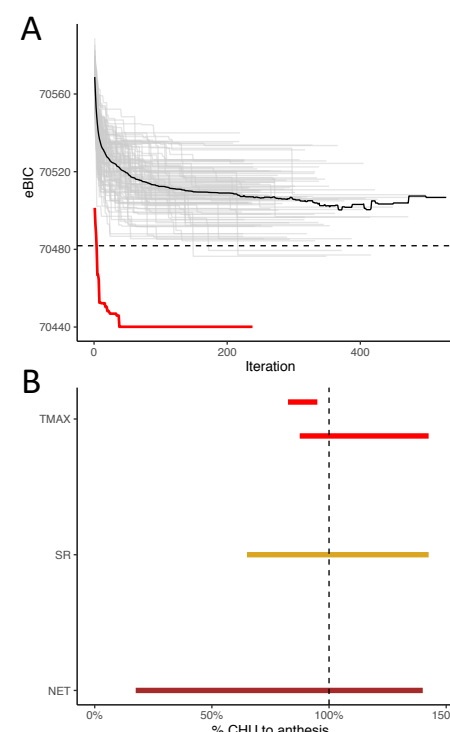
A. Selection of environmental variables was performed using a genetic algorithm with the extended BIC as the fitness measure (gray lines). Commonly selected variables were used in a second run to further prune variables (red line) and produce a final model (dashed line). **B.** Maximum daily temperature, incident solar radiation, and net evapotranspiration were the most important variables. The time over which they are important relative to male flowering (100% CHU to anthesis) is shown. **C.** The responses of 727 maize hybrids to these variables were calculated using random regressions and used as input for GWAS.
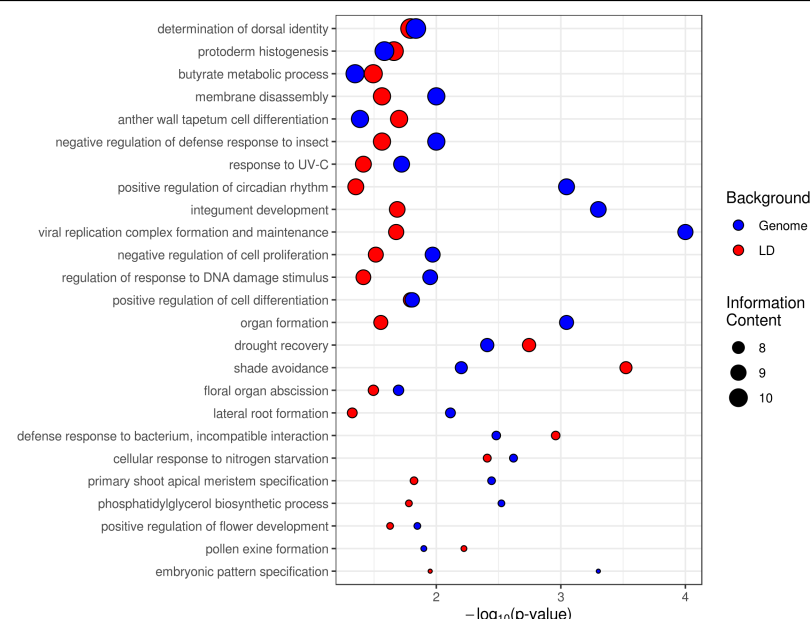
GO term enrichment was performed for candidate genes from GWAS. The 25 most informative terms are shown in the figure. GWAS identified genes that tend to play roles in root and flower development, drought and UV-C responses, and the development of the protective integument and pollen exine layers. This validates our selection of environmental variables and provides hypotheses for further research into the molecular mechanisms of plastic responses.
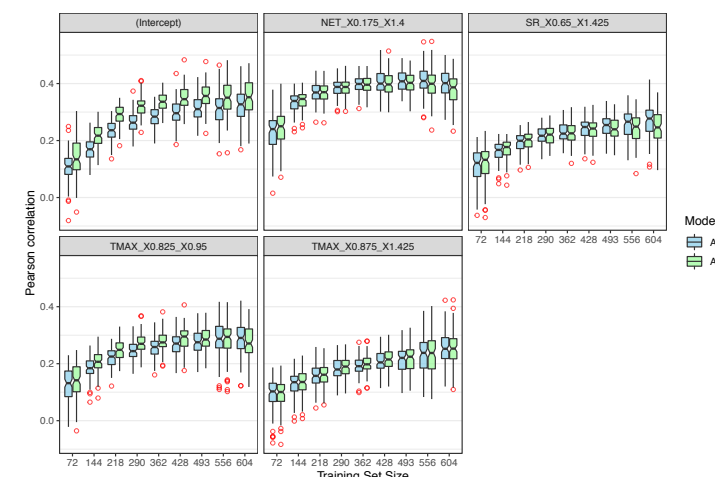
We also developed genomic prediction models for each plastic response and demonstrated moderate prediction accuracy for all responses. Thus, our method identifies environmental variables that can also be used to build genomic prediction models for use in breeding.